

Erfassung kognitiver Aspekte des Antwortverhaltens in einer Delphi-Studie

Häder, Michael; Rexroth, Margrit

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Häder, M., & Rexroth, M. (1998). *Erfassung kognitiver Aspekte des Antwortverhaltens in einer Delphi-Studie*. (ZUMA-Arbeitsbericht, 1998/06). Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-200528>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

ZUMA-Arbeitsbericht 98/06

**Erfassung kognitiver Aspekte des Antwortverhaltens
in einer Delphi-Studie**

Michael Häder, Margrit Rexroth

September 1998

ZUMA

Quadrat B2,1

Postfach 12 21 55

68072 Mannheim

Telefon: Häder: (0621) 12 46 – 136

Rexroth: (0621)12 46 - 230

Telefax: (0621) 12 46 - 100

E-mail: haeder@zuma-mannheim.de

rexroth@zuma-mannheim.de

Gliederung

1. Einleitung	3
2. Zur Diskussion um die Delphi-Methode	3
2.1 Generelle Diskussionen	3
2.2 Die Bildung der Expertenurteile	4
2.3 Zur Orientierung der Experten in den einzelnen Wellen	6
3. Möglichkeiten und Grenzen kognitiver Interviews	7
4. Design der Zusatzstudie	8
5. Die Wertigkeit der Problemlösung bei Experten und Nichtexperten	10
5.1 Die Lösung von Verständnis-, Abstraktions- und Wissens- Problemen	11
5.1.1 Nichtexperten	11
5.1.2 Experten	12
5.1.3 Zusammenfassung	14
5.2 Die Lösung des Prognose- und des Selbstreflexions-Problems	14
5.2.1 Nichtexperten	14
5.2.2 Experten	15
5.2.3 Zusammenfassung	18
6. Die Schätzungen in der zweiten Welle, quantitative Ergebnisse	19
7. Die Schätzungen in der zweiten Welle, qualitative Ergebnisse	22
7.1 Ergebnisse der kognitiven Interviews	23
7.2 Analyse der Befragungsergebnisse	24
7.2.1 Annäherung zum oder Entfernung vom Gruppenmittel	24
7.2.2 Stabilität und angegebene Kompetenz	26
8. Zusammenfassung	28
Literatur	30
Anlagenverzeichnis	32

1. Einleitung

In diesem Bericht¹ werden vier Fragen untersucht:

1. Sind Expertenschätzungen in einer Delphi-Studie höherwertig als die Schätzungen von Nichtexperten?
2. Inwieweit orientieren sich bei Delphi-Studien Experten und Nichtexperten an den Rückmeldungen der Ergebnisse aus der vorangegangenen Welle?
3. Sind die Schätzungen in der zweiten Runde einer Delphi-Befragung höherwertig als die Schätzungen in der ersten Runde?
4. Inwieweit ist es möglich, diese Fragen (1. bis 3.) im Rahmen einer speziellen Zusatzstudie, bei der gezielte Nachfragen eingesetzt werden, zu beantworten?

Zunächst wird der Stand der Diskussion um die Delphi-Methode kurz zusammengefaßt (Abschnitt 2). Vor diesem Hintergrund werden die Möglichkeiten und Grenzen kognitiver Techniken betrachtet (Abschnitt 3) und das Design einer von uns durchgeführten Zusatzstudie beschrieben (Abschnitt 4). Die Aufmerksamkeit gilt dabei vor allem den kognitiven Konstruktionsprinzipien, welche die beteiligten Personen bei ihrer Urteilsbildung benutzen. Mit Hilfe der Gegenüberstellung von Experten und Nichtexperten wird versucht zu zeigen, wodurch sich Expertenurteile auszeichnen (Abschnitt 5). Es wird danach analysiert (Abschnitt 6), inwieweit sich Experten und Nichtexperten bei ihren Schätzungen in der Folgerunde an den Rückmeldungen orientieren. Schließlich gilt die Aufmerksamkeit der Frage, ob sich mittels kognitiver Befragungen auch eine höhere Qualität der Schätzungen in der zweiten Welle nachweisen läßt (Abschnitt 7). Insbesondere sollen Hinweise auf die Wirkung von Konformitätsdruck und von zusätzlich aktiviertem Wissen gewonnen werden. Zusammenfassend (Abschnitt 8) versprechen wir uns Hinweise sowohl auf die im Zusammenhang mit der Delphi-Methode diskutierten Probleme als auch Aufschlüsse methodischer Art zu den Möglichkeiten der benutzen kognitiven Verfahren.

2. Zur Diskussion um die Delphi-Methode

2.1 Generelle Diskussionen

Grundsätzliche Zweifel am Funktionieren des Delphi-Ansatzes finden sich in der neueren Literatur nicht mehr (vgl. Häder/Häder 1998). Vor allem die inzwischen wahrgenommene Zu-

¹ Unser Dank gilt Frau Dr. Angelika Glöckner-Rist und Frau Professor Nanny Wermuth für ihre Hinweise zu diesem Manuskript.

verlässigkeit von Delphi-Prognosen und erfolgreiche Anwendungen bei sehr unterschiedlichen Problemstellungen sind hierfür wahrscheinlich wichtige Gründe. Diese mittlerweile nachweisbaren Erfolge sind allerdings auch darauf zurückzuführen, daß der Delphi-Methode zunächst von verschiedenen Kritikern mit starken Zweifeln begegnet wurde, die zum Teil bis zur generellen Ablehnung dieser Methode reichten. Die in diesem Zusammenhang angemahnten methodischen Forschungen und Evaluationsbemühungen und die zu dieser Thematik geführte Kontroverse haben die Entwicklung von Delphi jedoch insgesamt positiv beeinflußt.

Die umfassendste Kritik an der Delphi-Methode stammt wahrscheinlich von Sackman (1974, eine ausführliche Replik legen Strauß/Zeigler 1975:28 vor). In seinem Aufsatz fordert er die Anwender von Delphi auf, professionelle Standards zu entwickeln und Forschung zur Funktionsweise von Delphi zu betreiben, um die empirischen Ergebnisse solcher Studien abzusichern (vgl. Häder/Häder 1998).

In dieser Arbeit wird versucht, kognitionspsychologische Modelle und Methoden (Probing-Verfahren) für eine Erfassung mentaler Aspekte des Antwortverhaltens in Delphi-Studien nutzbar zu machen. Wenn es sich dadurch zeigen ließe, *wie* es Experten gelingt, *richtige Urteile* (z.B. Prognosen) über Sachverhalte zu fällen, über die ihnen zunächst nur unvollständiges Wissen vorliegt, so hätte dies für die Entwicklung und Beurteilung konkreter Delphi-Designs wesentliche Implikationen. Auch die bisher teilweise sehr unterschiedlichen Erfahrungen bei der Anwendung und Evaluation von Delphi könnten sich durch den Einsatz solcher Techniken erklären lassen. Damit kann schließlich sogar ein wesentlicher Teil der gegenüber der Delphi-Methode vorgetragenen Kritik zurückgewiesen werden.

Schließlich ist auf Diskussionen zu verweisen, die einen praktischen Aspekt betreffen. Es geht um die Frage, wodurch sich Experten überhaupt von Nichtexperten unterscheiden und damit darum, wer als Experte für eine Delphi-Studie gewonnen werden sollte.

2.2 Die Bildung der Expertenurteile

Zur Beantwortung der Frage, warum angenommen werden kann, daß mehrere Wiederholungen der Urteilsabgabe durch Experten bei Delphi zu qualitativ höherwertigen Schätzungen führen ziehen Häder/Häder (1995) kognitionspsychologische Ansätze und Ergebnisse heran. Dabei gehen die Autoren davon aus, daß es sich bei den Expertenschätzungen innerhalb einer Delphi-Studie um Urteilen unter Unsicherheit handele. Bei solchen Urteilen "hängt der Erfolg entscheidend davon ab, wie gut Informationssuche, Informationsbewertung und Informationsverknüpfung gelingen" (Zimolong/Rohrmann 1988:625).

Zur Erklärung der Bildung der Expertenurteile in der ersten Runde wird das *Informationsverarbeitungsparadigma* herangezogen. Danach können generell - also auch in Alltagssituationen - unvollständige Informationen aus dem Wissensbestand des Individuums ergänzt werden. Dazu nehmen die Individuen keine passiv-photographische Abbildung der Umwelt vor, sondern konstruieren diese intern aktiv. Schließlich kommt es zur zyklischen Rückkopplung zwischen Wahrnehmung, Gedächtnis und Informationssuche.

Die *Theorienklasse mentaler Modelle* bietet einen Zugang zum näheren Verständnis solcher Urteile. Für den Erfolg einer Schätzung bei unsicherem Wissen ist entscheidend, ob ein lokales mentales Modell (lokales MM) oder ein probabilistisches mentales Modell (PMM) genutzt wird. Während bei lokalen MM nur direkt auf die Aufgabe bezogenes Wissen aktiviert wird, wird bei der Konstruktion von PMM in größerem Umfang Erfahrungswissen aus dem Alltag (oder Expertenwissen) herangezogen (vgl. Gigerenzer et al. 1991).

In den Urteilen der folgenden Wellen spielen Lernprozesse, wie sie etwa Zimolong/Rohrman (1988:628) beschreiben, aber auch die Rückmeldungen selbst eine jeweils spezifische Rolle. So stellt Schwarz (1991, vgl. auch Hippler/Schwarz/Noelle-Neumann 1989, Schwarz/Sudman 1992 und Schwarz/Strack/Hippler 1990) auf der Grundlage von Arbeiten zur flexiblen Konstruktion mentaler Repräsentationen (Barsalou 1987, 1989) zur Urteilsbildung innerhalb einer Befragung fest, daß die Untersuchungspersonen auf verschiedene Arten von Wissen zurückgreifen. Bestimmtes Wissen sei den Befragten permanent verfügbar, andere Informationen dagegen nur temporär. In diesem Zusammenhang ist wichtig, daß (zunächst) nicht alle potentiell relevanten Informationen aus dem Gedächtnis abgerufen werden, sondern daß der Suchprozeß abgebrochen wird, sobald die Befragten genügend Informationen erinnert haben, um sich mit hinreichender Sicherheit ein Urteil bilden zu können (Bodenhausen/Wyer 1987). "Das Urteil beruht daher nur auf der Teilmenge potentiell relevanter Information, die zum Urteilszeitpunkt verfügbar ist - sei dies chronisch oder temporär" (Schwarz 1991:72). Die den Teilnehmern einer Delphi-Befragung nach jeder Welle zurückgemeldeten Informationen können schließlich bewirken, daß diese den "Suchprozeß" in ihrem Gedächtnis nochmals aufnehmen und nach weiteren, für den jeweiligen Sachverhalt relevanten Informationen fahnden. Die rückgemeldete Gruppenantwort kann einen Kontexteffekt bewirken und damit zur Verbesserung des abgegebenen Urteils beitragen. Zugleich ist die Rückmeldung aber auch als eine neue Information zu interpretieren. Sie muß in das bei den Experten bereits bestehende mentale Modell eingepaßt oder abgewiesen werden. Dazu müssen die Befragten prüfen, ob die neue Information in bezug auf das bestehende Modell als widerspruchsfrei integrierbar oder als abzuweisen eingeschätzt wird.

2.3 Zur Orientierung der Experten in den einzelnen Wellen

Die kognitiven Aspekte des Urteilsverhaltens der Experten in den einzelnen Wellen werden von verschiedenen Autoren untersucht. Insbesondere die Frage, welche Prozesse eine Meinungsänderung der Experten in die *richtige* Richtung bewirkt, stellt eine Herausforderung dar. So wird unterschieden zwischen den "Swingern", die ihre Antworten in einer Folgewelle ändern, da sie beispielsweise in ihrer Haltung unsicher sind und (oder) in ihren Urteilen am weitesten vom wahren Wert entfernt liegen, und jenen Experten, die am besten geschätzt haben und die ihre Antworten nicht ändern, da sie sich in ihren Einschätzungen sicher(er) fühlen. Das Funktionieren von Delphi setzt damit voraus, daß die Experten über Wissen bzw. zumindest über ein gewisses Gefühl zur Sicherheit ihres Urteils verfügen (Rowe 1991:238).

Einige Kritiker betrachten jedoch die Meinungsänderungen der Experten im Verlauf der Delphi-Befragung lediglich als Tendenz zum Gruppenmittel und nicht als eine Annäherung an den vermeintlich wahren Wert.

Der Beziehung zwischen der Stärke der Dissonanz der Experten gegenüber der Gruppe und der Tendenz zur Meinungsänderung waren - unter Rückgriff auf die Literatur zur Einstellungsänderung - bereits Experimente gewidmet. Ausgangspunkt ist die Überlegung, daß die Rückmeldung als extern gesetzter kognitiver Anker fungiert, der die Urteilsbildung beeinflussen kann. Es existieren drei Möglichkeiten:

1. der Anker wird ignoriert oder auch verarbeitet, trotzdem bleibt das Urteil letztlich unverändert,
2. das Urteil wird vom Anker weg verändert, z.B. um die Gesamtmeinung näher an die eigene zu bringen (Kontrasthaltung) oder
3. die Meinung wird zum Anker hin verändert (Assimilationshaltung). Es wird vermutet, daß je größer die Dissonanz mit der Gruppe ist, desto stärker auch der Druck zur Assimilation ausfällt.

Nach Aronson et al. (1963) ist die Meinungsänderung vor allem eine Funktion der Glaubwürdigkeit der anderen Experten und der Größe des eigenen Abstandes vom Gruppenergebnis. Diese Annahme wurde bei Nichtexperten, überprüft. Diese mußten im Rahmen einer Delphi-Befragung weniger glaubwürdige Urteile abgeben. Das Ergebnis zeigte zunächst, daß es bei Teilnehmern mit einem nur geringen Abstand vom Mittelwert ebenfalls eine Tendenz zur Konformität gab. Jedoch ergab sich weiterhin, daß diese Tendenz bei Teilnehmern, die in ihren Urteilen einen größeren Abstand zur Gruppenmeinung hatten, nicht mehr auftrat. Dies spricht für

die Hypothese, daß eine mangelnde Glaubwürdigkeit der anderen Testteilnehmer dazu beiträgt, daß hier keine Tendenz zum Konsens auftritt (Bardecki 1987:282ff.).

3. Möglichkeiten und Grenzen kognitiver Techniken zur Erfassung des Antwortverhaltens

Um den bei der Beantwortung von Fragen ablaufenden kognitiven Prozeß transparenter zu machen kann man mit verschiedenen kognitiven Techniken arbeiten (vgl. Sudman/Bradburn/Schwarz 1996:18ff.). Für den Einsatz dieser Techniken geht man davon aus, daß Personen in der Lage sind, solche Überlegungen zu verbalisieren, die sie ihren Antworten zugrunde legen. In bezug auf den Prozeß der Beantwortung von Fragen werden verschiedene kognitive Phasen unterschieden: Erstens muß die Frage verstanden und interpretiert werden (comprehension), danach wird zweitens im Gedächtnis nach relevanter Information gesucht (retrieval), drittens wird eine Antwort erstellt (judgement), die dann viertens noch zu editieren ist (response).

Eine Technik zur Erfassung kognitiver Aspekte des Antwortverhaltens stellt das auf Schuman (1966) und Belson (1981) zurückgehende Probing dar. Unter Probing versteht man Nachfrage-techniken, die das Ziel verfolgen, von den Befragten über die eigentliche Antwort hinausgehende, zusätzliche Information zu erhalten. Dabei kann es sich sowohl um Probings zum Fragenverständnis (Comprehension Probing) als auch um Nachfragen zu den Aspekten der Informationsbeschaffung (Information Retrieval Probing) handeln. Je nach Zeitpunkt der Nachfragen spricht man von "Follow-Up-Probings" (Nachfragen unmittelbar nach der Beantwortung der Frage) und "Post-Interview-Probings" (Nachfragen nach Beendigung des gesamten Interviews).

Vor- und Nachteile dieser Techniken nennen Prüfer/Rexroth (1996:109; vgl. auch Someren/Barnard/Sandberg 1994; Sudman/Bradburn/Schwarz 1996). Ein Nachteil von Nachfrage-techniken kann in einer a posteriori Rationalisierung der Entscheidungsfindung durch die Teilnehmer liegen, d.h. wenn diese ihre Erinnerung an die Problemlösung an deren Resultat orientieren. "When ask for memories, people answer a question not from direct memory of the cognitive process but from an interpretation that can easily be influenced by expectations." (Someren/Barnard/Sandberg 1994:21). Außerdem kann nicht ausgeschlossen werden, daß die Nachfragen das Problemlösungsverhalten selbst beeinflussen.

Die Lösung bestimmter Aufgaben bzw. die Beantwortung einer Frage führt in der Regel zu einem objektivierbaren Ergebnis. Aus kognitionspsychologischer Sicht ist es interessant, auf welche Weise und mit welchem Zeitaufwand eine befragte Person ein Ergebnis erzielt. Auch für die Entwicklung von Befragungsinstrumenten liefern Pretests zur Erfassung kognitiver Aspekte

des Antwortverhaltens wertvolle Hinweise. Eine andere Situation liegt bei Delphi-Studien vor. Hier lässt sich zunächst kein objektivierbares bzw. richtiges Ergebnis bestimmen. Trotzdem haben wir uns für den Einsatz kognitiver Techniken entschieden. Unser Interesse gilt der Beobachtung, welche unterschiedlichen Denkopoperationen, Modelle oder Wissensquellen von den verschiedenen Teilnehmern einer Delphi-Befragung mobilisiert werden, um zu einer Schätzung zu gelangen.

4. Design der Zusatzstudie

Von den Autoren wurde eine spezielle Zusatzstudie durchgeführt. Diese Zusatzstudie war ergänzender Bestandteil einer bei ZUMA betreuten Delphi-Befragung zum Thema ‘Demographische Veränderungen - Zukünftige Anforderungen und Strategien im Handwerk’ (im weiteren kurz: ”Zukunft des Handwerks”).² Der Titel der Delphi-Studie beschreibt bereits die inhaltliche Ausrichtung: Im Rahmen einer Expertenbefragung sollten u.a. detaillierte Voraussagen für das Jahr 2010 zu den Gebieten Arbeitsangebot, Arbeitsnachfrage, Arbeitszeitgestaltung und Tarifrecht im Handwerk gefunden werden. Die Befragung wurde in zwei Wellen 1997/1998 umgesetzt. Befragt wurden von den Projektnehmern³ ausgesuchte 387 Experten mit einem schriftlichen, standardisierten Instrument. Die inhaltlichen Ergebnisse, detaillierte Prognosen zum Handwerk, werden von den Projektnehmern selbst ausgearbeitet und an anderer Stelle präsentiert, sie werden hier nicht weiter behandelt. Im vorliegenden Bericht geht es ausschließlich um die Darstellung und die Ergebnisse der Zusatzstudie (vgl. Häder/Rexroth 1997, 1998).

Dieser Zusatzstudie verfolgte eine primär methodische Zielsetzung, die in der Weiterentwicklung der Delphi-Methode besteht. Ziel ist es, die vier eingangs genannten Probleme (vgl. Abschnitt 1) zu bearbeiten.

Dieser Zusatzstudie fand im Raum Mannheim, Heidelberg und Speyer statt. Die erste Welle wurde im Mai 1997 und die zweite im Februar 1998 erhoben. Befragt wurden insgesamt einmal zehn und einmal elf Personen. Diese wurden in der ersten Welle auch um die Beantwortung einiger demographischer Fragen gebeten. Alle Interviews wurden auf Tonband aufgezeichnet.

² Gefördert wurde diese Delphi-Studie vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie im Rahmen des Schwerpunktes ‘Demographischer Wandel und die Zukunft der Erwerbsarbeit im Standort Deutschland’

³ Die Projektnehmer sind folgende drei Verbundinstitute: die Zukunftswerkstatt e.V. der Handwerkskammer Hamburg, die Fachhochschule Niederrhein in Mönchengladbach und das Institut für sozialwissenschaftliche Forschung in München. Die Koordination des Projekts wurde von der Zukunftswerkstatt e.V. der Handwerkskammer Hamburg übernommen.

Das Frageprogramm bestand jeweils aus einem vierseitigen, von den Teilnehmern im Beisein der Interviewer auszufüllenden Fragebogen. Die Fragen waren Teil der Delphi-Studie zur "Zukunft des Handwerks". Zusätzlich wurden die Befragten in beiden Wellen mit einem kognitiven Instrument konfrontiert. Dieses enthielt gezielte Nachfragen zur Erfassung mentaler Prozesse bei der Beantwortung des präsentierten Fragebogens aus der Delphi-Studie.

Einbezogen wurden sechs Experten, und für Vergleichszwecke vier Nichtexperten. Bei den Experten handelt es sich um Mitarbeiter verschiedener Handwerksorganisationen. Wir gehen davon aus, daß diese Personen aufgrund ihrer beruflichen Tätigkeit dazu in der Lage sind, mittelfristige Entwicklungen im Handwerk zu beurteilen. Bei den Nichtexperten handelt es sich um Handwerker, die lediglich jeweils in einer bestimmten Handwerksbranche tätig sind und die nicht über Detailkenntnisse zum Handwerk insgesamt verfügen. (Die Testteilnehmer werden im Anhang I näher beschrieben und die Instrumente werden im Anhang III vorgestellt.)

Durch die gezielten Nachfragen in der Zusatzuntersuchung sollen in Anlehnung an Oksenberg/Cannell/Kalton (1991) drei Arten Verständnisprobleme aufgedeckt werden:

1. Die Bedeutung bestimmter Begriffe des Fragetextes soll erläutert und interpretiert werden,
2. Aspekte der Antwort sollen erklärt werden und
3. die Verständlichkeit von Begriffen wird hinterfragt, d.h. wie verständlich ist ein Begriff bzw. welche Probleme treten dabei auf.

Zusätzlich wurden die am Test beteiligten Personen um Confidence Ratings gebeten, mit deren Hilfe der Grad an Verlässlichkeit der eigenen Antwort ermittelt werden sollte. Wegen der relativ geringen Anzahl befragter Zielpersonen werden die Ergebnisse lediglich deskriptiv analysiert.

In der Delphi-Befragung mußten von den Teilnehmern verschiedene Aufgaben explizit und implizit bewältigt werden. Sie lassen sich *fünf Problem-Klassen* zuordnen:

- Verständnis-Problem,
d.h. das Verständnis der Aufgabenstellung dieser Delphi-Befragung. Aufgrund von spontanen Rückfragen der Teilnehmer sowie vermittelt über gezielte Nachfragen durch die Interviewer lassen sich Unklarheiten und Mißverständnisse über das Anliegen der Studie erkennen.
- Abstraktions-Problem
d.h. die Fähigkeit zur Abstraktion vom Bereich "Handwerk". Erklärtes Ziel des Projekts "Zukunft des Handwerks" waren globale Aussagen zur Situation im Jahr 2010. Dies erfor-

derte von den beteiligten Personen eine hohe Abstraktionsleistung. Zur Bewältigung dieser Aufgabe wurde in der ersten Welle durch die Interviewer eine explizite Nachfrage gestellt.

- **Wissens-Problem**

Die Zusatzstudie schloß in der ersten Welle Wissensfragen zu verschiedenen Begriffen ein. Erfragt wurde die Bedeutung der Termini: ‘Demographischer Wandel’, ‘Altersteilzeit’, ‘Organisationsgrad der Beschäftigten’ und ‘Organisationsgrad der Unternehmer in den Arbeitgeberverbänden’.

- **Prognose-Problem**

Schließlich waren 36 Prognosenaufgaben⁴ zu bearbeiten, d.h. es waren 36 Items unter dem Aspekt der Entwicklung des Handwerks bis zum Jahr 2010 zu bewerten.

- **Selbstreflektions-Problem**

Die beteiligten Personen wurden durch die Interviewer um Auskunft über die Quellen des Antwortverhalten sowie über die Sicherheit der abgegebenen Urteile gebeten. Dazu dienten neben zwei Indikatoren in den von den Teilnehmern auszufüllenden Fragebögen diverse Nachfragen durch die Interviewer in beiden Befragungswellen.

5. Die Wertigkeit der Problemlösung bei Experten und Nichtexperten

In diesem Abschnitt soll der Frage nachgegangen werden, ob sich eine höhere Qualität der Expertenschätzungen gegenüber den Urteilen der Nichtexperten nachweisen läßt und worauf dies gegebenenfalls zurückzuführen ist. Im Abschnitt 5.1 werden das Verständnis-, das Abstraktions- und das Wissens-Problem betrachtet und danach (Abschnitt 5.2) das Prognose- und das Selbstreflexions-Problem. Vor allem interessieren die Erkenntnishilfen (‘cues’), welche von den Befragten bei der Beantwortung herangezogen werden.

5.1 Die Lösung von Verständnis-, Abstraktions- und Wissens-Problemen

5.1.1 Nichtexperten

In dieser Gruppe traten spontan Probleme beim Verständnis der gesamten Aufgabenstellung und bei der Beantwortung der Fragen auf. Obwohl versucht wurde, die Aufgabenstellung verständlich zu formulieren, fühlten sich einige Nichtexperten überfordert und baten um Erklärun-

⁴ Da bei zwei Fragestellungen eine wesentliche Veränderung im Fragetext zwischen den beiden Wellen vorgenommen wurde, beschränkt sich die weitere Auswertung auf die verbleibenden 34 Indikatoren.

gen. Von diesen Befragten wurde direkt darauf verwiesen, daß die Beantwortung der Fragen schwer falle.

- **Verständnis-Problem**

Teilweise wurde das mit der Delphi-Studie verfolgte Ziel nicht richtig verstanden bzw. konnte von der Zielperson nicht bearbeitet werden. Der folgende Gesprächsausschnitt illustriert nicht zuletzt auch ein mangelndes Verständnis der Zielperson für das Ziel und den Inhalt der Untersuchung:

Interviewer: "Wie sicher waren Sie sich mit den einzelnen Urteilen, also bei den Fragen 'A' bis 'F'?"

Zielperson (ID07 - erste Welle): "Also eher unsicher."

Interviewer: "Und warum?"

Zielperson: "Weil ich mich mit der Materie eigentlich nicht in dem Maße auseinandergesetzt habe, weil das nicht zu meinem täglichen Leben gehört. Ich bin Handwerker, ich bekomme einen Auftrag und führe ihn aus, ich berechne ihn und beschäftige mich nicht damit, ob im Jahre 2010 ein Mädchen an irgendeinem Arbeitsplatz stehen wird, oder mehr Jugendliche Abitur machen oder Realschulabschluß."

- **Abstraktions-Problem**

Verallgemeinerungen, die Aufgabe, sich "das Handwerk insgesamt" vorzustellen, gelangen kaum bzw. nicht, wie in folgendem Beispiel:

Interviewer: "Woran haben Sie gedacht, als Sie eine Einschätzung zum Handwerk gegeben haben, haben Sie sich das Handwerk global vorgestellt, oder haben Sie sich einen Zweig im Handwerk gedacht?"

Zielperson (ID04 – erste Welle): "Ja ich hab' an mein Handwerk gedacht, Zentralheizungs- und Lüftungsbau. Ich kann keine Prognose für das Fleischerhandwerk geben."

In den Fällen, in denen verallgemeinernde Überlegungen zum Handwerk angestellt wurden, resultierten sie vor allem aus der Kommunikation mit anderen Handwerkern.

Interviewer: "Haben Sie bei den einzelnen Antworten eher geraten oder haben Sie ganz konkrete Hinweise zur Verfügung gehabt?"

Zielperson (ID05 – erste Welle): "Konkrete Hinweise jetzt nicht, aber ..."

Interviewer: "Was waren das denn für Hinweise?"

Zielperson (ID05): "Die man selbst auf der Baustelle, im Büro, bei den Kunden Gespräche auffängt."

Interviewer: "Haben Sie dabei an das Handwerk global gedacht oder an bestimmte Handwerkszweige?"

Zielperson (ID05): "Jetzt an bestimmte Handwerkszweige und zwar an Heizung und Sanitär."

- **Wissens-Problem**

Die nachgefragten Begriffe waren den Nichtexperten mitunter unklar und/oder wurden hinterfragt. Schwierigkeiten bereiteten vor allem die Ausdrücke "Demographischer Wandel" (obwohl dieser Begriff schon zu Beginn der Befragung erklärt wurde) und "Organisationsgrad der Beschäftigten". Mit "Altersteilzeit" wurden mehr oder weniger konkrete Inhalte verbunden. Probleme traten auch beim Ausdruck "Ungleicher Arbeitsanfall" auf.

5.1.2 Experten

- **Verständnis-Problem**

Das Ausfüllen des Fragebogens und das Verständnis der Aufgabenstellung verursachte kaum Probleme. Die Fragen erschienen als ganz normal:

Zielperson (ID11 – erste Welle): "Das sind eigentlich ganz normale Fragen, auf die man auch eine feste Antwort geben kann."

- **Abstraktions-Problem**

Auch für Verallgemeinerungen zum "Handwerk insgesamt" standen den Experten Denkstrategien zur Verfügung. Zwei unterschiedliche Wege haben wir angetroffen:

Erstens wurde - wie auch bereits bei den Nichtexperten - aufgrund von eigenen Erfahrungen mit unterschiedlichen Handwerksbereichen ein verallgemeinerndes Urteil über das Handwerk insgesamt abgeleitet. Da die Experten jedoch tätigkeitsbedingt über ein größeres berufliches

Umfeld verfügten, flossen in die Einschätzungen mehr Erfahrungen ein und diese fielen somit differenzierter aus als bei den Nichtexperten.

Zweitens wurden von den Experten Verallgemeinerungsstrategien eingesetzt. Beispielsweise wurden Erfahrungen aus *typischen* Handwerksbereichen herangezogen. Die Situation von Randgruppen wurde bewußt weniger in den Mittelpunkt gestellt (wie sich bei den folgenden Gesprächen zeigt).

Interviewer: "Es ging um Handwerk. Haben Sie da an das Handwerk allgemein gedacht oder haben Sie sich bestimmte Bereiche vorgestellt?"

Zielperson (ID11 - erste Welle): "Wenn man solche Fragen gestellt bekommt, dann denkt man in erster Linie an die Bereiche, die im Handwerk besonders stark vertreten sind. Ich hab' das vorhin schon genannt, einmal der Bau als Marktführer oder als Konjunkturindikator und zum zweiten der KFZ-, Metallbereich, wo wir auch eine ganze Reihe von Zulieferbetrieben auch haben, die abhängig sind von ihrem Abnehmer, und an die denkt man in erster Linie. Man denkt weniger an Randgruppen, zum Beispiel Damen- oder Herrenschneider, die da auch noch existieren und soweit sie es verstehen auch ganz gut existieren, aber von der Zahl der Betriebe her eigentlich unbedeutend sind. Also Metallbereich, das ist bei uns mit Abstand der größte Bereich von den Betrieben her. 42 bis 43% Baubereich und innerhalb des Metallbereichs der Sonderbereich KFZ."

Interviewer: "Und bei der Frage 2.2. spricht man ja nun allgemein vom Handwerk, haben Sie sich da das Handwerk global vorgestellt oder haben Sie an bestimmte Bereiche gedacht?"

Zielperson (ID01 - erste Welle): "Gut, sich es global vorzustellen ist ein bißchen schwierig, weil's eben inzwischen noch 130 Handwerksberufe gibt. Einige von denen, da wird allerdings schon gar nicht mehr ausgebildet. Nein da denke ich schon eigentlich an einen Metallbau, Ausbau, an die großen Gruppen."

Bei den folgenden Antworten wurden dann teilweise differenzierte Auskünfte gegeben, die jeweils nur für bestimmte Bereiche des Handwerks Geltung besitzen sollten. So wurden wieder-

holt die Besonderheiten des Bauhandwerks bei bestimmten Arbeitszeitregelungen hervorheben.

- Wissens-Problem

Die erfragten Begriffe waren allen Experten geläufig und wurden von allen in ihrer Bedeutung richtig verstanden.

5.1.3 Zusammenfassung

Zwischen beiden Gruppen bestehen die erwarteten Unterschiede. Insgesamt erwiesen die Experten bei ihren Antworten zu den Wissens-, Abstraktions- und Verständnis-Problemen eine deutlich größere Kompetenz. Außerdem wurde die Aufgabenstellung kaum als außergewöhnlich betrachtet. Anders in der Gruppe der Nichtexperten. Hier waren die einzelnen Personen bei der Beantwortung der Fragestellungen überfordert. Für sie war die Art der Aufgabenstellung außergewöhnlich. Obwohl einige Nichtexperten überfordert waren und dies auch explizit zum Ausdruck brachten, wurde der schriftliche Fragebogen von allen Beteiligten fast vollständig ausgefüllt.

Dieser Befund unterstreicht nicht zuletzt auch die Notwendigkeit und Bedeutung einer sorgfältigen Expertenauswahl bei Delphi-Studien. Die gezeigten Ergebnisse können dazu beitragen, das Expertendenken zu erkennen und somit Kriterien zu finden, die für die Rekrutierung der Experten benutzt werden können.

5.2 Die Lösung des Prognose- und des Selbstreflexions-Problems

5.2.1 Nichtexperten

- Prognose-Problem

Bei den Antworten spielten drei Aspekte eine Rolle: Es wurde *erstens* angegeben, sich bei den Antworten auf den eigenen *direkten Erlebnisbereich* zu stützen, d.h. es wurde auf Erfahrungen aus dem Berufsleben zurückgegriffen. Die zu bewertenden Szenarien wurden *zweitens* vor allem aufgrund ihrer *Machbarkeit* im eigenen (Handwerks-)Betrieb beurteilt. *Drittens* wurde die *Entwicklung* der letzten Jahre - soweit bekannt - *fortgeschrieben*. Hier wahrgenommene Trends wie der wachsende Anteil ausländischer Arbeitskräfte wurden für die Zukunft verallgemeinert. Dazu die folgenden Beispiele, wobei der zuerst zitierte Text zugleich auch einen Hinweis auf das Verständnis-Problem enthält:

Zielperson (ID07 – erste Welle): "Immer wenn für einen Handwerker was aus der Praxis ist, dann ist es einfacher zu beantworten als wenn es hochgestochen formuliert ist, daß man erst eine längere Zeit darüber nachdenken muß, weil es eben nicht direkt in den Bereich fällt, mit dem ein Handwerker tagtäglich zu tun hat."

Interviewer: "Kennen Sie Prognosen, die Sie für Ihre Antworten auf diese Fragen genutzt haben?"

Zielperson (ID04 – erste Welle): "Nein."

Interviewer: "Haben Sie bei Ihren Antworten dann eher geraten oder antworteten Sie aufgrund mehr oder weniger konkreten Wissens?"

Zielperson (ID04): "Wieder gefühlsmäßig."

Interviewer: "Und was meinen Sie damit?"

Zielperson (ID04): "So wie ich den Eindruck habe, daß es sich entwickelt."

- **Selbstreflexions-Problem**

Das Reflektieren über das eigene Antwortverhalten fiel bei den Nichtexperten relativ undifferenziert aus. Der folgende Gesprächsausschnitt erscheint uns dafür typisch:

Interviewer: "Haben Sie bei Ihren Antworten eher geraten oder haben Sie auf konkretes Wissen hin die Fragen beantwortet?"

Zielperson (ID07 – erste Welle): "Ich muß sagen, daß ich das eigentlich sehr gefühlsmäßig mache. Ich versuche die Frage zu verstehen und beantworte sie dann so, wie ich meine, daß es abläuft, ohne fundamentiertes Wissen."

5.2.2 Experten

Die Expertengruppe war heterogener zusammengesetzt als die der Nichtexperten. So trafen bei den einzelnen Teilnehmern jeweils nur bestimmte der im folgenden vorgestellten Aspekte zu.

- **Prognose-Problem**

Erstens lassen sich einige formale Ähnlichkeiten zur Gruppe der Nichtexperten erkennen. Mit Hilfe von *Extrapolationen* wurden Schätzungen zu künftigen Trends vorgenommen. Die

Experten verfügten über umfassende Kenntnisse zu den bisherigen Entwicklungen. Sie beschäftigten sich teilweise beruflich mit der Erstellung entsprechender Bilanzen (Konjunkturstatistik für die Handwerkskammer (ID01), regionale Arbeitsmarktstudien für die Industrie- und Handelskammer (ID09) bzw. verfügten über relevantes Datenmaterial zu diesem Aspekt (ID11)). Auch hierzu ein Beispiel:

Zielperson (ID09 - zweite Welle): "Bei der Frage 1.1 nahm ich an, daß es weiterhin sehr wahrscheinlich ist, daß wir ein Arbeitskräfteüberangebot haben. Einfach weil man die Aktualität fort schreiten kann, weil ich nicht davon ausgehe, daß der EURO zu einer erheblichen Entlastung des Arbeitsmarktes beiträgt. Hier habe ich mich durch die 61%⁵ bestätigt gesehen. Das hat aber meine Entscheidung nicht all zu stark beeinflußt."

Es kam aber auch vor, daß die Grenzen der bisherigen Entwicklung erkannt wurden, wie in der folgenden Stellungnahme zu einer anderen Prognoseaufgabe:

Zielperson (ID08 - erste Welle): "Ich glaube nicht, daß die Entwicklung wie in den letzten Jahren weitergeht, daß die Geringfügigkeitsgrenze immer nach oben geschraubt wird. Ich bin schon lang' davon überzeugt, daß die 610,- Mark, die wir zur Zeit haben, irgendwo eine Grenzlinie darstellen, weil die 610,- auch in einem anderen beitragsrechtlichen Bereich vor Jahren schon eingefroren wurden. Die Grenze in der Vergangenheit, wer unter 610,- Mark verdient hat, versicherungspflichtig geworden ist, da hat der Arbeitnehmer beide Anteile zahlen müssen, ohne Abzug für den Arbeitnehmer, jetzt ist die Grenze erreicht, wer also in Zukunft könnte, würde also so weiter geht geringfügig sein, ... oder sagen wir es umgekehrt, jeder der heute eine Mark verdient kriegt gleich abgezogen. Damit ist das etwas aus den Fugen geraten. Ich denke also, daß die 610 Mark eingefroren wird, wenn es in den Geringfügigkeitsbereich hineingeht."

Zweitens spielte der *subjektive Erlebnisbereich* für die Urteilsfindung eine wichtige Rolle. Auch dieses Hilfsmittel stand den Experten hochwertiger zur Verfügung. So wurde in der Regel

⁵ Diese Zahl betrifft das aus der ersten Befragungswelle rückgemeldete Ergebnis.

auf eine langjährige Berufserfahrung z.B. in Dachorganisationen verwiesen, welche zugleich den Umgang mit einer Vielzahl von Handwerksbetrieben einschloß.

Drittens wurden auch die antizipierten *Realisierungschancen*, ähnlich wie bereits bei den Nichtexperten, bei den Prognosen mit einbezogen. Hier spielten die Kosten eine wesentliche Rolle:

Zielperson (ID08 - erste Welle): "Frage 'F' denke ich, wird in sehr starkem Umfang eine Kostenfrage sein. Wenn es finanzierbar ist: Ja; wenn es nicht finanzierbar ist: Nein. Aber ich denke, es ist nicht finanzierbar."

Die Experten zeichneten sich jedoch dadurch aus, daß sie über weitere kognitive Erkenntnis-hilfen verfügten. Dabei handelte es sich um die folgenden:

Viertens: Aufgrund von *Analogieschlüssen* wurden Urteile gefällt. Die Entwicklungen, die bereits in der Industrie erkennbar sind, würden zukünftig auch im Handwerk eintreten oder das Modell USA werde sich auch in der Bundesrepublik durchsetzen:

Zielperson (ID01- zweite Welle): "Sonntagsarbeit - der Dienstleistungsgedanke – wird zunehmen. Denke ich doch. Da hoffe ich, daß wir irgendwann einmal amerikanische Verhältnisse kriegen. Was nicht heißen soll, daß amerikanische Verhältnisse in allen Bereichen gut sind. In dem Bereich aber denke ich doch schon. Wie man ja auch sieht im Bäckerhandwerk als Beispiel."

Fünftens, es kam zur *Verallgemeinerung* oder zur *Übertragung* der Folgen von Entwicklungen in anderen Bereichen auf das Handwerk. Technische Trends und allgemeine gesellschaftliche oder politische Entwicklungen, wie die europäische Integration, spielten eine Rolle. Genannt wurden auch Folgen, die aus der Einführung der Computertechnik für den Handwerksbereich resultieren.

Sechstens wurden auf der Grundlage von *Leitbildern* bzw. *Ideologien* Schätzungen abgegeben. Beispielsweise gelte das Handwerk als konservativ und reagiere entsprechend fortschritts-feindlich, es würde sich jedoch der Flexibilität als "Gebot der Stunde" stellen müssen, der Dienstleistungsaspekt präge das Handwerk, moderne technologische Entwicklungen bestimmten die Entwicklung im Handwerk usw. Hierfür sollen wieder zwei Gesprächsausschnitte zitiert werden:

Interviewer: "Haben Sie bei den einzelnen Einschätzungen geraten, oder haben Sie da konkrete Hinweise zur Verfügung, an denen Sie sich orientiert haben?"

Zielperson (ID09 – erste Welle): "Konkretes Wissen nicht, aber es ist natürlich ein in mir steckendes Leitbild, das von einer zunehmenden Flexibilisierung in der Wirtschaft ausgeht. Da sind viele Fragen die mit diesem Leitbild automatisch eine gewisse Antwort implizieren."

Zielperson (ID11 – erste Welle): "Die Frage ist wirklich schwer zu beantworten, wie das Handwerk reagieren wird. Das Handwerk reagiert auf solche Entwicklungen immer sehr sehr zurückhaltend. Das Handwerk ist ja irgendwo erzkonservativ und deswegen müsste man hier sehr vorsichtig sein ..."

Siebtens, galten die gegenwärtig von jungen Handwerksmeistern geführten Betriebe als *Muster bzw. Vorausgruppe* für die Handwerksbetriebe der Zukunft. Hier momentan unternommene Bemühungen bei der Unternehmensführung wurden für die Zukunft des Handwerks insgesamt als typisch interpretiert.

Achtens wurde in einigen wenigen Fällen auch auf *direkte Erfahrungen* mit Prognosen bzw. auf relevantes Wissen zur weiteren Entwicklung des Handwerks verwiesen. In diesem Zusammenhang wurden eigene Umfragen und Literaturquellen genannt. Im Sinne der Theorienklasse mentaler Modelle (Gigerenzer et al. 1991) haben wir es an dieser Stelle mit lokalen MM zu tun, bzw. - im Verständnis des Modells der Informationsverarbeitung in Befragungssituationen - liegt eine Art von "Recall prior judgement" (Sudman/Bradburn/Schwarz 1996:58) vor, wie wir im Abschnitt 2.2 bereits dargestellt haben.

- Selbstreflexions-Problem

Die Schwierigkeit von Prognosen und eigene Unsicherheiten bei den Antworten wurden auch von den Experten artikuliert. Falls es den Experten nicht möglich war, ausreichend kognitive Quellen für ihre Antworten zu erschließen, wurden auch von ihnen *spontane Auskünfte* gegeben. Dies erfolgte jedoch insgesamt seltener als bei den Nichtexperten.

5.2.3 Zusammenfassung

Von den Nichtexperten wurde auf die entsprechende Nachfrage angegeben, vor allem nach dem Gefühl geantwortet zu haben. Prognosen über die Bevölkerungsentwicklung waren in der Regel nicht bekannt. Niemand aus der Gruppe der Nichtexperten hatte sich bisher mit der in der

Befragung aufgegriffenen Prognose-Problematik befaßt. Je abstrakter der zu beurteilende Sachverhalt war, desto schwieriger fiel den Nichtexperten eine Antwort. Die beschriebenen Denkstrategien (viertens bis achtens) können dagegen als typische Expertendenkweisen angesehen werden. Während die Nichtexperten mit dem Thema der Befragung überfordert waren und demzufolge kaum verwertbare Antworten gegeben haben dürften, waren die Aufgaben von den Experten prinzipiell lösbar. Auch bei den einzelnen Experten haben wir unterschiedliche Denkmuster zur Problemlösung angetroffen. Als wesentlich stellte sich der individuelle Erlebnisbereich heraus. Darüber hinaus kann vermutet werden, daß einige Experten mit Hilfe der beschriebenen Hilfsmittel offenbar dazu in der Lage sind, ein fundierteres Urteil zu bilden. Sie zeichnen sich damit durch zwei Aspekte aus: ihnen steht kognitiv ein umfangreicheres Datenmaterial zur Verfügung und die darauf aufbauenden Synthesen erfolgen aufgrund komplizierterer Denkprozesse.

Die kognitiven Interviews führten damit zu einem eindeutigen Befund in der erwarteten Richtung.

6. Die Schätzungen in der zweiten Welle, quantitative Ergebnisse

Die in der zweiten Befragungswelle rückgemeldeten Ergebnisse⁶ erlauben es, eine zweite Problematik zu behandeln: Auf welche Weise orientieren sich bei Delphi-Studien Experten – im Unterschied zu Nichtexperten - an den Rückmeldungen der Ergebnisse aus der vorangegangenen Welle?

Die gezielten Nachfragen zeigten zunächst, daß die Rückmeldungen von den verschiedenen Teilnehmern jeweils unterschiedlich zur Kenntnis genommen wurden⁷. Das folgende Beispiel zeigt, wie eine Zielperson mit den rückgemeldeten Ergebnissen umgeht:

⁶ Es handelt sich um das zusammengefaßte Ergebnis der ersten Welle der von der Zukunftswerkstatt e.V. der Handwerkskammer Hamburg im Rahmen des Projekts "Zukunft des Handwerks" koordinierten Expertenbefragung (n= 387).

⁷ Bei den Antworten spielten - nicht zuletzt aufgrund des relativ langen Zeitabstandes von ca. neun Monaten - die Erinnerungen an die erste Befragungswelle keine Rolle.

Zielperson: (ID 01 – zweite Welle). "Wie groß war Ihre Stichprobe?"

Interviewer: "387 Experten aus dem Handwerk, aus der Wissenschaft und aus sonstigen Institutionen."

Zielperson (ID 01, während des Beantwortens von Frage 1.3 'A'): "Da werde ich mit der Mehrheit gehen, das halte ich für 'eher wahrscheinlich' und für absolut wünschenswert. Das ist gar keine Frage." ...

Interviewer: "Sie haben bei der Frage 1.3 'E', dort angekreuzt, wo die meisten denken, daß es nicht so sein wird. Spielte hier die Mehrheit keine Rolle für Sie?"

Zielperson (ID 01): "Gerade bei 'E' denke ich schon, daß gerade wenn wir in den Bereich der Dienstleistungen kommen, wo auch mehr den Bereichen Marketing mit allem was dazu kommt Beachtung geschenkt wird, da denke ich, daß gerade eine Teilzeitbeschäftigung für qualifizierte weibliche ... egal, ob das jetzt in die Richtung Akquisition geht, selbst wenn es in die Richtung Mahnwesen geht, daß man einfach hinterher telefonieren kann, daß das mehr bringt, daß da was passieren wird. Deshalb halte ich es für sehr wahrscheinlich." ...

Interviewer: "Spielten die Rückmeldungen bei Ihren Einschätzungen eine Rolle? Haben Sie sich daran orientiert?"

Zielperson(ID 01): "Eigentlich kaum."

Interviewer: "Ja, warum eigentlich nicht? Waren Sie sich so sicher?"

Zielperson(ID 01): "Aus meiner Sicht eigentlich ja."

Tabelle 1 zeigt, in welcher Relation die erneuten Urteile der Teilnehmer nun tatsächlich zur Rückmeldung bzw. zur ersten Schätzung stehen.

Tabelle 1: Anteil der Konstraturteile, der Assimilationen und der unveränderten Schätzungen bei allen 34 Variablen⁸, Angaben in Prozent

	n	Kontrast- urteile	Assimilationen	kein Unterschied	Missing	N (= 100%)
Alle,	10	16.5	25.6	54.4	3.5	340
Davon:						
Experten	6	13.2	24.5	58.3	3.9	204
Nichtexperten	4	21.3	27.2	48.5	2.9	136

Einer starken Assimilationshaltung, d.h. einer - von Delphi-Kritikern aufgrund eines normativen Gruppendrucks befürchteten - *ausschließlichen* Orientierung an diesen Informationen sind wir nicht begegnet. In mehreren Fällen behaupteten die Zielpersonen sogar, die Rückmeldungen völlig ignoriert zu haben. Damit wären eigentlich noch mehr Kontrasturteile zu erwarten gewesen.

Die Experten unterscheiden sich hinsichtlich der Assimilation faktisch nicht von den Nichtexperten. Die Experten gaben aber deutlich weniger häufig Kontrasturteile ab und urteilten stattdessen stabiler als die Nichtexperten. Dies könnte darin begründet sein, daß für die Experten die Rückmeldungen eine höhere Autorität darstellen und sie diese stärker zur eigenen Orientierung benutzen als die Nichtexperten. Auch können konstante Antworten auf eine stärkere Auseinandersetzung mit der Thematik hindeuten, die dann zu überlegteren Urteilen führt. Es zeigte sich immerhin, daß es in beiden Gruppen (Experten und Nichtexperten) häufiger zu Assimilationsurteilen als zu Kontrasturteilen kam.

Ein solches Ergebnis war nicht zu erwarten. Plausibel wäre eher das entgegengesetzte Resultat: Die in ihren Urteilen weniger sicheren Nichtexperten orientieren sich stärker an den Rückmeldungen als die kompetenteren Experten. (Es wurde gegenüber den Befragten vor der zweiten Befragung betont, daß es sich um die Rückmeldung der Schätzungen von 387 "tatsächlichen" Experten handelt.) Eine Erklärung für dieses Phänomen läßt sich nicht direkt aus den Ergebnissen der kognitiven Interviews ableiten. Indirekt wurde jedoch deutlich, daß die Nichtexperten die Befragung mit einer nur geringen Ernsthaftigkeit absolviert haben (Abschnitt 5.1). Auch waren die einzelnen Fragen für sie zu schwierig zu beantworten. Für das Ziel der Befragung wurde von den Nichtexperten außerdem weniger Verständnis aufgebracht. Damit ist es plausibel, daß für die Nichtexperten das rückgemeldete Ergebnis der ersten Welle – wie auch die gesamte Befra-

⁸ Im Anhang II (vgl. Tabelle A6) werden die vollständigen Ergebnisse gezeigt.

gung - nur von relativ geringem Interesse war. Demgegenüber, so könnte man vermuten, haben sich die stärker an den inhaltlich Ergebnissen interessierten Experten intensiver mit ihren eigenen und den rückgemeldeten Antworten befaßt. Allerdings sind die Fallzahlen zu gering, um hier Verallgemeinerungen vornehmen zu können.

Schließlich haben auch psychologische Experimente gezeigt, daß eigenes Wissen im Allgemeinen überschätzt wird. So wurde festgestellt, daß Urteile von denen die Testteilnehmer annahmen, sie seien zu 100 Prozent richtig, nur zu 80 Prozent zutrafen (Hoffrage 1993:7). Offenbar zeichnen sich Experten auch dadurch aus, daß sie in dieser Beziehung in ihren Urteilen ebenfalls sicherer sind.

Die weiteren Analysen (Abschnitt 7.) zur Meinungsänderung in der zweiten Welle stützen sich aufgrund der offenbar geringen Qualität der Nichtexpertenurteile nur auf die Schätzungen der Experten.

7. Die Schätzungen in der zweiten Welle, qualitative Ergebnisse

Eine dritte in unserer Studie bearbeitete Frage ist, ob die Schätzungen in der zweiten Welle höherwertig sind als die der ersten und damit eine Ergebnisverbesserung der Schätzungen eingetreten ist. Damit wird eine wichtige, die Legitimation des Delphi-Ansatzes stützende Annahme überprüft. Ein solcher Nachweis ist jedoch nicht ohne weiteres realisierbar, da eine zuverlässige Evaluation der Schätzergebnisse vor Ablauf des Prognosezeitraums im Jahr 2010 nicht möglich ist. Deshalb muß nach indirekten Anhaltspunkten gesucht werden.

Der Delphi-Ansatz geht von der Überlegung aus, daß die Experten sich bei den wiederholten Schätzungen mit den ihnen rückgemeldeten Ergebnissen auseinandersetzen (vgl. Häder/Häder 1998). Zwei Möglichkeiten sind denkbar:

1. Im Falle einer von den einzelnen Experten stärker vermuteten eigenen Kompetenz werden die ursprünglichen Urteile reproduziert und die Rückmeldung gegebenenfalls ignoriert. Es wird vorausgesetzt, daß die Experten auch über adäquates Wissen zur eigenen Kompetenz verfügen.
2. Bei Unsicherheit über das eigene Urteil wird bei der wiederholten Schätzung stärker dem Gruppenurteil gefolgt.

(In beiden Fällen wird angenommen, daß den anderen an der Delphi-Studie beteiligten Experten ausreichend Kompetenz zugesprochen wird und die Rückmeldungen entsprechend bewertet werden.)

Nachdem gezeigt werden konnte (Abschnitt 5.), daß die Experten bei ihren ersten Schätzungen, zumindest im Vergleich zu den Nichtexperten, kompetente Urteile abgegeben haben, würde ein solches Vorgehen bei der zweiten Schätzung indirekt für eine Verbesserung der Gruppenleistung sprechen. Aufgrund der Nachfragen in den kognitiven Tests sowie aufgrund der vorliegenden Antworten wird es möglich, diese Vermutungen zu überprüfen.

7.1 Ergebnisse der kognitiven Interviews

Die Resultate der kognitiven Test sind überraschend ausgefallen: Fast alle Experten stritten ab, sich bei ihren erneuten Antworten an der Rückmeldung orientiert zu haben. Dazu einige typische Beispiele:

Interviewer: "Als Sie die Fragen beantwortet haben, woran haben Sie sich da primär orientiert? Haben Sie es gefühlsmäßig ausgefüllt, haben Sie es auf der Grundlage konkreten Wissens beantwortet, haben Sie danach geschaut, was die anderen Teilnehmer geantwortet haben oder haben Sie sich vielleicht sogar erinnert, was Sie vor circa einem Jahr geantwortet hatten?"

Zielperson (ID11 – zweite Welle): "Ich habe jetzt fast nicht auf die Prozentzahlen geguckt. Ich habe mich auch nicht an den Mehrheitszahlen orientiert". ...

Interviewer: "Gibt es noch eine weitere Vorgabe, bei der Sie bewußt von der Meinung der anderen abweichen?"

Zielperson (ID11): "Ich habe da gar nicht so sehr auf die Prozentzahlen geguckt, sondern ich habe das bewußt jetzt auf der Seite liegen lassen, bin aber in beiden Fragen wohl bei der Mehrheit gelandet. Aber nicht, weil ich jetzt meinte, mich der Mehrheit anschließen zu müssen. Daß möchte ich jetzt besonders betonen, ich habe versucht, das auszuklammern."

Interviewer. "Wie sehr haben Sie sich an den Rückmeldungen der anderen Teilnehmer orientiert?"

Zielperson (ID08 – zweite Welle): "Ich habe versucht, mich nicht von anderen irritieren zu lassen."

Interviewer: "Jetzt habe ich ein paar Fragen an Sie. Als Sie jetzt den Fragebogen ausgefüllt haben, haben Sie das eher auf der Grundlage Ihres Wissens gemacht, haben Sie es gefühlsmäßig gemacht oder haben Sie sich auch an den Antworten der anderen Teilnehmer, die wir hier hineingeschrieben haben, orientiert?"

Zielperson (ID03 – zweite Welle): "In keinsten Weise."

Interviewer: "Gab es auch irgendwelche Positionen, bei denen Sie sagten, da habe ich mich bewußt entgegen der Meinung der anderen entschieden? Oder haben Sie diese Meinungen gar nicht interessiert?"

Zielperson (ID03): "Das habe ich gar nicht beachtet"

Diese Kommentare geben zunächst Grund zu der Annahme, daß es in der zweiten Befragungswelle zu keiner Ergebnisverbesserung gekommen ist. Die Rückmeldungen wurden offenbar ignoriert bzw. bei der Suche nach einer erneuten Antwort nicht einbezogen und stattdessen lediglich eine nochmalige, davon relativ unabhängige Schätzung abgegeben. Dies widerspräche einer wichtigen Grundannahme für Delphi-Studien. Weiteren Aufschluß soll die Analyse der tatsächlich abgegebenen Schätzungen geben.

7.2 Analyse der Befragungsergebnisse

7.2.1 Annäherung zum oder Entfernung vom Gruppenmittel

Bei der Auswertung der Antworten zu den gezielten Nachfragen (Abschnitt 7.1) hatte sich ergeben, daß die Experten eine direkte Orientierung an den Rückmeldungen bestreiten. Dies spricht zunächst nicht für eine Verbesserung der Schätzungen, da es der Grundannahme des Erfolgs von Delphi-Studien widerspricht. Im weiteren wird deshalb beleuchtet, ob es nicht vielleicht doch zu einer, eventuell unbewußten, Auseinandersetzung mit der Gruppenmeinung gekommen ist. Insbesondere im Falle eines besonders großen Abstandes der eigenen ersten Schätzung vom Gruppenmittel wäre dies zu erwarten. Bei einer solchen Konstellation wäre theoretisch eine Orientierung an diesen Rückmeldungen zu vermuten. Es wird also nach einen Zusammenhang gesucht zwischen erstens dem Abstand der ersten Schätzung von der Gruppenmeinung und zweitens der Stärke des Bestrebens, sich bei der wiederholten Schätzung dieser Gruppenmeinung anzunähern. Da eine Erinnerung an das erste Urteil aufgrund des großen zeitlichen

Abstandes zwischen beiden Schätzungen ausgeschlossen werden kann, wäre eine solche unbewußte Orientierung durchaus plausibel.

Geht man davon aus, daß ein besonders großer Abstand zur Gruppenmeinung zu einer erneuten kognitiven Auseinandersetzung mit der Schätzaufgabe anregt, dann würde eine Anpassung an die Gruppenmeinung für eine Verbesserung der Urteile sprechen. Ließe sich jedoch kein Zusammenhang zwischen der ursprünglichen Größe des Abstands vom Gruppenniveau und der Tendenz finden, sich diesem anzunähern, so müßte man von spontanen und willkürlichen, die Gruppenmeinung ignorierenden und schließlich den Delphi-Ansatz infragestellenden Schätzungen berichten.

Die Ergebnisse der Gegenüberstellung von erstens dem Abstand der ersten Schätzung zu dem bei der zweiten Schätzung rückgemeldeten Gruppenmittel und zweitens dem Abstand der zweiten Schätzung von diesem Gruppenmittel sind in der Graphik 1⁹ sowie in den Kreuztabellierungen im Anhang II enthalten.

Wir sind hier offenbar an eine Grenze der benutzten kognitiven Methode gestoßen: Die Experten sind nicht dazu in der Lage, aufgrund unserer Nachfragen über die Rückmeldungen als Quellen für ihr Antwortverhalten zu berichten. Bei einer nichtreaktiven Analyse der Antworten zeigt sich ein deutlich anderes Bild als bei den Nachfragen durch die Interviewer. Dies ist zugleich ein Hinweis darauf, daß der bei Delphi-Studien vorausgesetzte Mechanismus funktioniert.

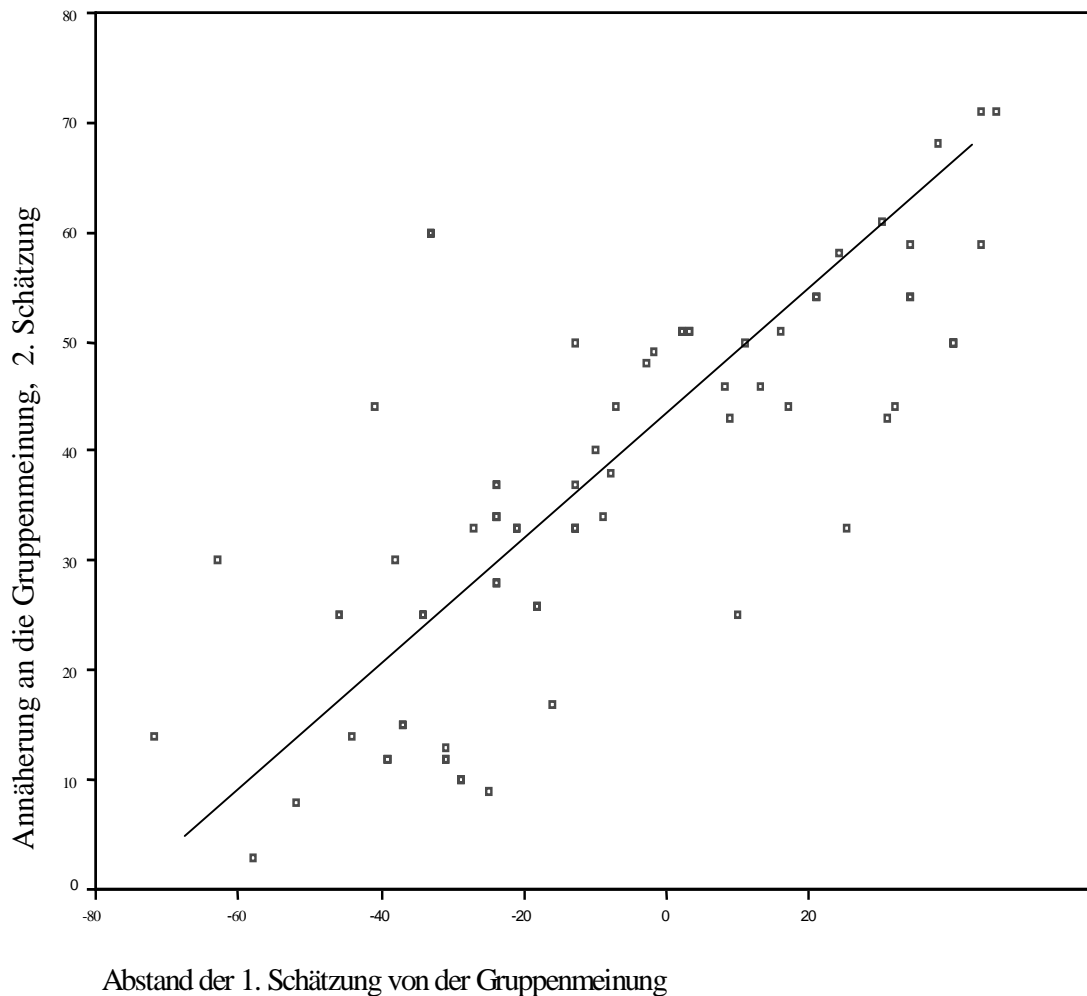
Bei der Graphik läßt sich eine Tendenz zu einer monoton steigenden Funktion feststellen (Pearsons Korrelationskoeffizient: .486, sign.: .000). Dies bedeutet, daß jene Teilnehmer, die in der ersten Welle eine Einschätzung abgegeben haben, die besonders weit von der Gruppenmeinung entfernt liegt, sich nun dieser am stärksten annähern während umgekehrt jene Teilnehmer, die bereits in der ersten Welle mit ihrem Urteil näher am mainstream gelegen haben, dieses nicht ändern bzw. jetzt eher zu Kontrasturteilen neigen. Hier handelt es sich nun um eine relativ eindeutige Ergebnislage.

⁹ Für die Bestimmung der Orientierung der Zielpersonen an der Rückmeldung wurde folgendes Vorgehen benutzt: Jeder Teilnehmer bekam für jede Antwort einen Punktwert zugeordnet. Dieser Wert resultiert aus der prozentualen Anzahl der Experten der Hauptbefragung (n=387) die ebenfalls genau diese Antwortvariante gewählt haben. Damit bedeutet eine relativ große Zahl, daß diese Meinung von vielen Experten geteilt wird bzw. eine relativ große Nähe zum Gruppenurteil.

Mit Hilfe dieser Werte läßt sich nun ein Vergleich zwischen erster und zweiter Welle vornehmen. Von einer Kontrasthaltung wird dann gesprochen, wenn die Differenz zwischen diesem Wert aus der ersten und aus der zweiten Welle positiv ist, negative Werte bedeuten ein Assimilationsurteil.

Es wurden in der Graphik lediglich Schätzungen berücksichtigt, bei denen in der zweiten Welle eine Veränderung stattgefunden hat.

Graphik 1: Scatterplot: Abstand der ersten Schätzung von der Gruppenmeinung und Stärke der Annäherung an diese Gruppenmeinung bei der zweiten Schätzung



Vergleicht man die Aussagen der Experten mit ihrem tatsächlichen Antwortverhalten, so zeigt sich ein deutlicher Widerspruch: Während behauptet wird, die Rückmeldungen nicht bzw. nicht sonderlich zur Kenntnis genommen zu haben, zeigt der Vergleich mit den tatsächlich abgegebenen Antworten aus der ersten Welle einen gegenteiligen Befund. Die Antworten der Experten bewegen sich merklich in Richtung der ihnen rückgemeldeten Ergebnisse. Dies geschieht um so stärker, je weiter sie in ihren ersten Schätzungen von der Gruppenmeinung entfernt gelegen haben.

7.2.2 Stabilität und angegebene Kompetenz

Die gezielten Nachfragen legen nahe, daß von den Experten die Rückmeldungen ignoriert worden sind (Abschnitt 7.1) und es eventuell zu keiner Verbesserung der Schätzungen gekommen ist. Im vorigen Abschnitt wurden nun die tatsächlich abgegebenen Schätzungen analysiert

und dabei jedoch ein anderes, für das Funktionieren von Delphi insgesamt positives, Ergebnis gewonnen (Abschnitt 7.2.1). Eine weitere, in diesem Zusammenhang durchaus naheliegende Möglichkeit besteht nun darin, die Experten nach der selbst zugeschriebenen Kompetenz zu gruppieren und danach die Stabilität der Antworten in den jeweiligen Gruppen zu untersuchen.

Eine Voraussetzung für das Funktionieren von Delphi bzw. für eine Verbesserung des Gruppenurteils bei den wiederholten Schätzungen ist ein mehr oder weniger sicheres Wissen der Experten über die Zuverlässigkeit der eigenen Schätzung bzw. über den Grad an eigener Kompetenz zum jeweiligen Thema. Weiter wird angenommen, daß bei einer stärker empfundenen eigenen Kompetenz die Rückmeldung weniger beachtet und stattdessen das eigene Urteil reproduziert wird. Fühlen sich die Experten bei den jeweiligen Themenbereichen jedoch nicht (so) kompetent, dann liegt eine stärkere Orientierung an den Schätzungen der anderen Experten bzw. an den Rückmeldungen nahe (Abschnitt 2.2). Dies könnte dann als Indiz für eine verbesserte Schätzung in der Folgewelle gewertet werden. Damit ist es angezeigt, an dieser Stelle auch diesen Zusammenhang zu überprüfen. Auskunft über die Stabilität der Urteile zwischen beiden Wellen bei den jeweiligen Fragekomplexen gibt zunächst Tabelle 2.

Tabelle 2: Veränderungen im Antwortverhalten: Durchschnittliche Anzahl der Kategorien, in denen sich die ersten von der zweiten Schätzung unterscheidet¹⁰

	Veränderungen im Antwortverhalten bei den Fragen ...				Summe der Veränderungen
	1.1 und 1.2	1.3, B - F	2.1, A – M	2.2, A – E	
Experten	0.7	3.3 ¹¹	4.7	4.5	13.2
Nichtexperten	1.8	7.3	7.8	4.8	21.5

In Tabelle 2 wird nochmals deutlich, daß die Experten stabilere Urteile abgegeben haben als die Nichtexperten. Dies wurde bereits als Hinweis auf die Zuverlässigkeit der Expertenschätzungen interpretiert, kann jedoch noch nicht als Indiz für eine Qualitätsverbesserung der Schätzungen gelten. Aus Tabelle 3 geht nun hervor, in welchem Maße sich die Experten bei den einzelnen Themenkomplexen Kompetenz zuschreiben.

¹⁰ Beim Übergang von einem 'Missing' zu einer gültigen Schätzung (und auch im umgekehrten Fall) wurde jeweils der Wert '1' codiert. Als Antwortkategorien werden die im Fragenbogen nebeneinander liegenden Vorgaben aufgefaßt.

¹¹ Die Zielperson ID11 übersah beim Ausfüllen des Fragebogens in der zweiten Welle eine komplette Spalte. Die so entstandenen 'Missing-Werte' werden jedoch nicht als eine Änderung des Antwortverhaltens interpretiert.

Tabelle 3: Anzahl der Experten (n=6), die sich im jeweiligen Themenbereich in der zweiten Welle für sehr, teilweise und gering kompetent halten

Kompetent zu:	Arbeitsangebot u. Arbeitsnachfrage			Arbeitszeitgestaltung u. Tarifrecht		
	sehr	teilweise	gering	sehr	teilweise	gering
Anzahl d. Experten	1	2	3	0	1	5

Die Darstellung zeigt, daß an dieser Stelle leider keine weiteren quantitativ-vergleichenden Auswertungen, differenziert nach dem von den Experten angegebenen Kompetenzgrad, vorgenommen werden können. Die einzelnen Zellen sind teilweise nicht bzw. nur schwach besetzt, sodaß sich hier solche Vergleiche nicht möglich sind. Es wird deshalb lediglich über einen gewissen Trend berichtet, ohne dabei Zahlen anzugeben:

Jene Experten, die sich selbst für kompetent(er) halten, neigen dazu, ihre Antworten weniger stark zu verändern. Wie beim Delphi-Ansatz erwartet, korrespondiert eine höhere Kompetenz mit stabileren Antworten. Der erwartete Zusammenhang geringere Kompetenz => stärkere Assimilation – kann jedoch nicht eindeutig festgestellt werden.

Als Ergebnis ist festzustellen, daß diese Analysen zu keiner eindeutigen Ergebnislage führen konnten. Zwar sind die Expertenurteile stabiler als die der Nichtexperten, jedoch lassen sich die erwarteten Zusammenhänge zwischen der selbst angegebener Kompetenz und der Antwortstabilität aufgrund einer zu geringen Besetzung der einzelnen Zellen nicht quantitativ untersuchen. Eindeutige Hinweise auf eine verbesserte Qualität der Schätzungen konnten so an dieser Stelle nicht gefunden werden.

8. Zusammenfassung

Vier Probleme standen im Mittelpunkt der Untersuchung:

Erstens war dies der Nachweis, daß die Expertenschätzungen innerhalb einer Delphi-Studie höherwertig sind als die Schätzungen von Nichtexperten. Ein Urteil darüber, wer im Rahmen des Tests eine richtige Schätzung abgegeben hat, kann nicht zuverlässig vor Ablauf der Prognosefrist getroffen werden. Im Vergleich zu den Nichtexperten haben die Experten jedoch kompetentere Urteile gebildet. Der von ihnen herangezogene Wissensvorrat ist wesentlich größer als bei den Nichtexperten. Schließlich war zu beobachten, daß einige Experten deutlich kreativer sind und komplexere Denkopoperationen ausführten, um zu einer Schätzung zu gelangen. Damit kann insgesamt angenommen werden, daß die Expertenschätzungen valider ausgefallen sind als die der übrigen Teilnehmer und daß schließlich auch aus den Ergebnissen der Delphi-Befragung

interessante Schlußfolgerungen für das Projekt "Zukunft des Handwerks" gezogen werden können.

Zweitens wurde untersucht, wieweit sich bei Delphi-Studien Experten und Nichtexperten an den Rückmeldungen der Ergebnisse aus der vorangegangenen Welle orientieren. Die Nichtexperten waren durch die Komplexität der Fragestellungen eindeutig überfordert und konnten dem Anliegen der Delphi-Studie nicht folgen. Sie waren deshalb auch nicht dazu motiviert, sich bei ihren Antworten in der zweiten Welle mit den Rückmeldungen auseinanderzusetzen. Das Kriterium "Nichtanpassung an die Gruppenmeinung" als Zeichen für besondere Kompetenz hat versagt. In dieser Studie wurde jedoch der von den Teilnehmern den Rückmeldungen beigemessene Grad an Glaubwürdigkeit nicht explizit ermittelt. Es wurde lediglich darauf verwiesen, daß 387 Experten aus verschiedenen Bereichen befragt worden sind (vgl. Anlage 3: Instrumente).

Drittens ging es um den schwierig zu erbringenden Nachweis, daß die Expertenschätzungen in der zweiten Runde einer Delphi-Befragung höherwertig sind als in der ersten Runde. Die zu erwartenden Zusammenhänge zwischen der verbal geäußerten Sicherheit bzw. der eigenen vermuteten Kompetenz einerseits und der Antwortstabilität andererseits zeigen sich zunächst – nicht zuletzt aufgrund der zu geringen Anzahl an Untersuchungspersonen - nur andeutungsweise. Nutzt man ein von der Selbstreflexion der Experten unabhängiges Kriterium, dann stellt sich der erwartete Zusammenhang jedoch relativ klar heraus. Immerhin liefern die Antwortverteilungen einen Beleg dafür, daß die Rückmeldungen von den Experten zur Kenntnis genommen worden sind und in die neuerliche Antwortfindung einbezogen wurden. Geht man von der besonderen Qualität der Experten-Schätzungen bereits in der ersten Welle aus, dann kann geschlossen werden, daß die Ergebnisse der zweiten Befragungsrunde von höherer Qualität sind.

Viertens war von Interesse, inwieweit solche Fragen mit Hilfe gezielter Nachfragen (Probes) beantwortet werden können. Die benutzten Instrumente haben zunächst gezeigt, worin sich die Expertenurteile von denen der Nichtexperten unterscheiden. Die gewählte Fragestrategie erbrachte jedoch nur unklare Hinweise darauf, wie die Experten bei den wiederholten Schätzungen zu ihren Urteilen gelangt sind. Die ursprünglich im Rahmen von Pretests zur Beurteilung des Erhebungsinstruments eingesetzten kognitiven Nachfragen müssen für Forschungen in diesem Zusammenhang weitaus differenzierter ausgearbeitet und gezielter den Fragestellungen angepaßt werden.

Literatur

- Arosen, E./ Turner, J./ Crasmer, J. (1963): Communication credibility and communication discrepancy as determinants of opinion change. In: *Journal of Abnormal Social Psychology* 67, S. 31-36.
- Bardecki, M.J. (1984): Participants' Response to the delphi Method: An Attitudinal Perspective. In: *Technological Forecasting and Social Change* 25, S. 281-292.
- Barsalou, L.W. (1987): The instability of graded structure: Implications for the nature of concepts. In: Neisser, U. (Hrsg.) *Concepts and conceptual development: Ecological and intellectual factors in categorization*, Cambridge University Press., S. 101-140.
- Barsalou, L.W. (1989): Intraconcept similarity and its implications for interconcept similarity. In: Vosniadou, S./ Ortony, A. (Hrsg.) *Similarity and analogical reasoning*, Cambridge: Cambridge University Press, S. 76-121.
- Belson, W.A. (1981): *The Design and Understanding of Survey Questions*. Aldershot, England: Gower.
- Bodenhorn, G.V./ Wyer, R.S. (1987): Social cognition and social reality: Information acquisition and use in the laboratory and the real world. In: Hippler, H.-J./ Schwarz, N./ Sudman, S. (Hrsg.) *Social information processing and survey methodology*, New York: Springer, S. 6-41.
- Gigerenzer, G./ Hoffrage, U./ Kleinbölting, H. (1991): Probabilistic Mental Models: A Brunswikian Theory of Confidence. In: *Psychological Review*, Vol. 98, No. 4, S. 506-528.
- Häder, M./ Rexroth, M. (1997): Beobachtungen im Rahmen einer Delphi-Studie: Fallbeschreibungen, 1. Welle, ZUMA-Technischer Bericht 97/09.
- Häder, M./ Rexroth, M. (1998) Beobachtungen im Rahmen einer Delphi-Studie: Fallbeschreibungen, 2. Welle, ZUMA-Technischer Bericht 98/04.
- Häder, M./ Häder, S. (1995): : Delphi und Kognitionspsychologie. Ein Zugang zur theoretischen Fundierung der Delphi-Methode. In: *ZUMA-Nachrichten* 37, Mannheim.
- Häder, M./ Häder, S. (1998): Neure Entwicklungen bei der Delphi-Methode, Literaturbericht II, ZUMA-Arbeitsbericht 98/05
- Hippler, H.-J./ Schwarz, N./ Noelle-Neumann, E. (1989): Response Order Effects in Dichotomous Questions: The Impact of Administration Mode. *ZUMA-Arbeitsbericht* 89/17. Mannheim.
- Hoffrage, U. (1993): Die Illusion der Sicherheit bei Entscheidungen unter Unsicherheit. In: Hell, W./ Fiedler, K./ Gigerenzer, G. (Hrsg.): *Kognitive Täuschungen. Fehl-Leistungen und Mechanismen des Urteilens, Denkens und Erinnerns*. Spektrum Akademischer Verlag: Heidelberg Berlin Oxford: 73-97.
- Oksenberg, L./ Cannell, Ch./ Kalton, G. (1991): New Strategies for Pretesting Survey Questions. In: *Journal of Official Statistics*, 7:349-365.
- Prüfer, P. /Rexroth, M. (1996): Verfahren zur Evaluation von Survey-Fragen: Ein Überblick. In: *ZUMA-Nachrichten* 39:95-115.
- Rowe, G./ Wright, G./ Bolger, F. (1991): Delphi. A reevaluation of Research and Theory: In: *Technological Forecasting and Social Change* 39, S. 235-251.

- Sackman, H. (1975): A Delphi Critique. Lexington Books, Lexington, Massachusetts.
- Schuman, H. (1966): The Random Probe: A Technique for Evaluation the Validity of Closed Questions. In: American Sociological Review, 31: 218-222.
- Schwarz, N. (1991): Assimilation und Kontrast in der Urteilsbildung: Implikationen für Fragereihenfolgeeffekte. In: ZUMA-Nachrichten 29, S. 70-86.
- Schwarz, N./ Sudman, S. (1992): Context Effects in social and psychological research. New York: Springer.
- Schwarz, N./ Strack, N./ Hippler, H.-J. (1990): Kognitionspsychologie und Umfrageforschung: Themen und Befunde eines interdisziplinären Forschungsgebiets. ZUMA-Arbeitsbericht 90/07. Mannheim.
- Somerén, M.W.van/ Barnard, Y.F./ Sandberg, J.A.C. (1994). The Think Aloud Method. A practical guide to modelling cognitive processes. Academic press: London Boston San Diego New York Sydney Tokyo.
- Sudman, S./ Bradburn, N.M./ Schwarz, N. (1996): Thinking about answers. The Application of Cognitive Processes to Survey Methodology. Jossey-Bass Publ. San Francisco.
- Strauss, H.J./ Zeigler, L.H. (1975): The Delphi Technique and Its Uses In Social Science Research. In: The Journal of Creative Behavior, vol 9, no 4.
- Zimmermann, B./ Rohrman, B. (1988): Entscheidungshilfetechnologie. In: Frey, D./ Graf Hoyer, C./ Stahlberg, D. (Hrsg.) Angewandte Psychologie. Weinheim: Psychologie Verlags Union.

Anlagenverzeichnis

1. Zusammensetzung der Befragengruppe

2. Tabellen:

Tabelle A1: Gesamtübersicht

Tabelle A2: Orientierung der Schätzungen an der Rückmeldung, Prozentsatz der Experten, die eine gleiche Schätzung abgegeben haben

Tabelle A3: Kreuztabellierungen Alternativfragen

Tabelle A4, Kreuztabellierungen, vierstufige Fragen

Tabelle A5, Kreuztabellierungen, dreistufige Fragen

Tabelle A6: Kontrasturteile, identische Schätzungen und Assimilation pro Frage bei allen Befragten, bei Experten und bei Nichtexperten

3. Instrumente:

- Schriftlich von den Experten in der ersten Welle zu beantwortender Fragebogen
- Mündliche Zusatzfragen an die Experten in der ersten Welle
- Schriftlich von den Experten in der zweiten Welle zu beantwortende Fragen einschließlich Rückmeldung der Ergebnisse der ersten Welle
- Mündliche Zusatzfragen in der zweiten Welle

Anhang I: Zusammensetzung der Befragtengruppe

Zur Gruppe der Experten innerhalb der gezielten Nachfragen zählten ursprünglich sieben Personen:

- ein Abteilungsleiter (ID02) und
- ein stellvertretender Geschäftsführer einer Innungskrankenkasse (ID08),
- ein Leiter eines Bildungs- und Technologiezentrums einer Handwerkskammer (ID03),
- zwei Unternehmensberater bei einer Handwerkskammer (ID01 und ID10)
- und ein Hauptgeschäftsführer einer Handwerkskammer (ID11) sowie
- ein Referent der Hauptgeschäftsführung einer Industrie- und Handelskammer (ID09).

Da ein Experte (ID10) den Mitschnitt seines Interviews verweigerte, stützt sich die Auswertung lediglich auf die übrigen sechs Experten. Dieser Experte wurde in der zweiten Welle nicht mehr befragt.

Die Gruppe der Nichtexperten setzt sich zusammen aus je einem selbständigen Handwerker in Kleinbetrieben in den Bereichen:

- Heizungsbau/Sanitär/Gas/Wasser (ID04),
- Zentralheizungs-/Lüftungsbau (ID05),
- Maler (ID06) und
- Fernmeldetechnik (ID07).

Anhang II: Tabellen

Tabelle A1: Gesamtübersicht

	ID:	1 ¹²	2	3	4	5	6	7	8	9	11
Ind.: 1	v1	2-2	3-3	2-1	3-3	3-3	4-2	2-2	3-3	3-3	2-2
1.1	v2	2-2	1-2	1-1	1-1	2-1	2-1	1-2	2-2	1-1	1-1
1.2	v3	2-2	2-1	1-1	2-1	3-2	1-2	2-1	3-2	1-1	3-2
1.3 B	v4	2-2	3-2	3-3	3-4	3-2	3-2	3-3	2-3	2-2	3-3
	v5	1-1	1-1	1-1	1-1	1-1	2-1	2-1	1-1	1-1	1--
C	v6	1-2s	1-3	2-1	1-2	1-1	2-1	4-3	2-3	3-2	1-2
	v7	1-1	1-1	1-1	1-1	1-1	1-1	2-1	1-1	1-1	1--
D	v8	2-3	3-2	1-2	2-3	1-4	1-1	1-3	3-3	2-2	3-3
	v9	1-1	1-1	1-1	1-2	2-2	1-1	1-2	1-1	1-1	---
E	v10	1-1	3-3	1-2	1-1	1-2	3-1	3-4	2-3	2-3	4-3
	v11	1-1	1-1	1-1	1-1	1-1	2-2	2-2	1-1	1-1	2--
F	v12	2-3	3-3	2-3	1-4	3-2	3-1	4-3	3-3	2-3	4-3
	v13	1-1	1-1	1-1	1-1	2-2	2-2	2-1	1-1	1-1	2--
2.	v14	2-3	3-3	2-3	2-3	2-1	3-2	2-3	3-3	4-3	2-2
2.1 A	v15	2-2	2-2	3-3	2-3	2-3	1-2	2-2	2-2	2-1	2-1
B	v16	1-1	1-1	2-3	2-2	3-1	1-2	2-1	2-1	1-1	3-1
C	v17	2-2	2-1	1-2	2-1	1-1	1-1	1-1	2-1	1-1	1-1
D	v18	2-3	1-2	1-1	1-2	2-2	1-3	1-1	2-3	2-2	--1
E	v19	1-1	1-2	2-1	2-2	1-3	2-1	1-1	1-2	1-1	1-2
F	v20	2-2	3-2	2-3	3-1	1-2	1-1	2-2	3-2	1-1	1-1
G	v21	2-2	1-1	1-1	1-1	1-1	1-3	1-1	1-1	1-1	1-1
H	v22	1-1	1-1	1-1	1-1	3-1	3-1	1-1	1-1	1-1	1-1
I	v23	1-1	1-1	1-1	1-2	2-1	1-1	1-1	1-2	1-1	2-1
K	v24	1-1	1-1	1-1	2-2	2-1	1-1	1-1	1-2	1-1	1-1
L	v25	1-3	2-3	3-3	2-2	3-2	--3	3-3	2-3	2-2	3-3
M	v26	2-2	2-2	2-2	2-3	2-2	--1	2-3	2-1	3-3	3-3
2.2 A	v27	3-3	2-2	2-2	2-3	4-4	4-3	3-3	2-3	2-2	1-1
	v28	1-2	2-2	1-1	1-2	2-2	2-1	2-2	1-2	2-1	1-1
B	v29	2-2	1-2	3-2	2-3	1-3	--3	1-1	1-2	3-2	2-2
	v30	1-1	1-2	2-1	2-1	1-1	--1	1-1	1-1	2-1	1-1
C	v31	2-1	1-2	2-2	2-1	1-1	2-1	1-1	1-2	2-2	1-2
	v32	1-1	1-1	1-1	1-1	1-1	1-1	1-1	1-1	1-1	1-1
D	v33	2-1	1-1	--2	1-1	1-2	1-3	1-1	1-2	2-1	1-2
	v34	1-1	1-1	--1	1-1	1-1	2-2	1-1	1-1	1-1	1-1
E	v35	2-1	2-2	3-1	3-2	1-1	1-3	1-1	1-2	1-1	1-2
	v36	1-1	1-1	2-1	2-1	1-1	2-2	1-1	1-1	1-1	1-1

¹² Die Angaben zu den Experten sind **fett** gedruckt, die der Nichtexperten normal.

Tabelle A2: Orientierung der Schätzungen an der Rückmeldung, Prozentsatz der Experten (n = 387), die eine gleiche Schätzung abgegeben haben

Fragen	1. Welle	2. Welle	Fragen	1. Welle	2. Welle	Fragen	1. Welle	2. Welle
1.1	49.0	49.0	2.1 A	47.8	38.8	2.2 A	30.5	37.4
1.2	34.5	43.0	B	38.8	49.4		49.5	49.2
1.3 B	43.6	38.0	C	52.6	61.8	B	30.2	41.1
	67.8	80.0	D	44.4	35.0		49.7	49.0
C	31.8	35.6	E	45.0	41.5	C	46.0	46.0
	81.1	89.0	F	34.5	40.7		94.0	94.0
D	26.7	33.9	G	57.0	52.0	D	37.7	39.2
	63.8	55.3	H	74.6	93.0		81.1	82.0
E	24.5	27.4	I	60.4	60.4	E	32.4	36.6
	60.5	64.8	K	62.2	62.2		64.4	78.8
F	28.6	35.3	L	43.9	53.8			
	55.8	67.9	M	45.3	38.6			

Tabelle A3: Kreuztabellierungen¹³ Alternativfragen

DRV28_1 * KONV28 Crosstabulation

Count

		KONV28			Total
		-3,00	,00	3,00	
DRV28_1	48,00	1	1		2
	51,00		2	2	4
Total		1	3	2	6

DRV30_1 * KONV30 Crosstabulation

Count

		KONV30			Total
		-2,00	,00	2,00	
DRV30_1	49,00	1	3		4
	51,00			2	2
Total		1	3	2	6

DRV36_1 * KONV36 Crosstabulation

Count

		KONV36		Total
		-72,00	,00	
DRV36_1	14,00	1		1
	86,00		5	5
Total		1	5	6

¹³ Erklärung der Variablen:

konv{x}: bei positiven Werten: Stärke des Kontrasturteils bzw. bei negativen Werten Stärke des Assimilationsurteils;

bei konstanten Urteilen ist der Wert 0, jeweils bei der Variable v{x}

drv{x}_1: Konformität mit der Rückmeldung bei der Variable {x} in der ersten Welle

Tabelle A4, Kreuztabellierungen, vierstufige Fragen

DRV2_1 * KONV2 Crosstabulation

Count

		KONV2		Total
		,00	30,00	
DRV2_1	31,00	2		2
	61,00	3	1	4
Total		5	1	6

DRV3_1 * KONV3 Crosstabulation

Count

		KONV3			Total
		-39,00	,00	16,00	
DRV3_1	12,00	2			2
	35,00		2		2
	51,00		1	1	2
Total		2	3	1	6

DRV4_1 * KONV4 Crosstabulation

Count

		KONV4			Total
		-8,00	,00	8,00	
DRV4_1	38,00	1	2		3
	46,00		2	1	3
Total		1	4	1	6

DRV6_1 * KONV6 Crosstabulation

Count

		KONV6					Total
		-44,00	-34,00	10,00	34,00	44,00	
DRV6_1	15,00	1					1
	25,00		2	1			3
	59,00				1	1	2
Total		1	2	1	1	1	6

DRV8_1 * KONV8 Crosstabulation

Count

		KONV8				Total
		-25,00	-9,00	,00	9,00	
DRV8_1	9,00	1				1
	34,00		1	1		2
	43,00			2	1	3
Total		1	1	3	1	6

DRV10_1 * KONV10 Crosstabulation

Count

		KONV10		Total
		-29,00	,00	
DRV10_1	10,00	2	1	3
	39,00		3	3
Total		2	4	6

DRV12_1 * KONV12 Crosstabulation

Count

		KONV12			Total
		-31,00	-18,00	,00	
DRV12_1	13,00	1			1
	26,00		3		3
	44,00			2	2
Total		1	3	2	6

DRV27_1 * KONV27 Crosstabulation

Count

		KONV27		Total
		-13,00	,00	
DRV27_1	6,00		1	1
	37,00	1	3	4
	50,00		1	1
Total		1	5	6

DRV29_1 * KONV29 Crosstabulation

Count

		KONV29			Total
		-37,00	-24,00	,00	
DRV29_1	15,00	2			2
	28,00		2		2
	52,00			2	2
Total		2	2	2	6

DRV31_1 * KONV31 Crosstabulation

Count

		KONV31			Total
		-24,00	,00	24,00	
DRV31_1	34,00	3			3
	58,00		2	1	3
Total		3	2	1	6

DRV33_1 * KONV33 Crosstabulation

Count

		KONV33			Total
		-21,00	,00	21,00	
DRV33_1	33,00	2	1		3
	54,00			2	2
Total		2	1	2	5

DRV35_1 * KONV35 Crosstabulation

Count

		KONV35				Total
		-16,00	-13,00	,00	13,00	
DRV35_1	17,00	1				1
	33,00		2	1		3
	46,00			1	1	2
Total		1	2	2	1	6

Tabelle A5, Kreuztabellierungen, dreistufige Fragen

DRV15_1 * KONV15 Crosstabulation

Count

		KONV15		Total
		,00	34,00	
DRV15_1	26,00	1		1
	54,00	3	2	5
Total		4	2	6

DRV16_1 * KONV16 Crosstabulation

Count

		KONV16				Total
		-52,00	-27,00	,00	25,00	
DRV16_1	8,00	1				1
	33,00		1		1	2
	60,00			3		3
Total		1	1	3	1	6

DRV17_1 * KONV17 Crosstabulation

Count

		KONV17			Total
		-46,00	,00	46,00	
DRV17_1	25,00	2	1		3
	71,00		2	1	3
Total		2	3	1	6

DRV18_1 * KONV18 Crosstabulation

Count

		KONV18			Total
		-10,00	,00	40,00	
DRV18_1	40,00	1	1		2
	50,00		1	2	3
Total		1	2	2	5

DRV20_1 * KONV20 Crosstabulation

Count

		KONV20			Total
		-31,00	,00	31,00	
DRV20_1	12,00	2			2
	43,00		1	1	2
	45,00		2		2
Total		2	3	1	6

DRV21_1 * KONV21 Crosstabulation

Count

		KONV21		Total
		-63,00	-33,00	
DRV21_1	30,00	1		1
	60,00		5	5
Total		1	5	6

DRV23_1 * KONV23 Crosstabulation

Count

		KONV23			Total
		-38,00	,00	38,00	
DRV23_1	30,00	1			1
	68,00		4	1	5
Total		1	4	1	6

DRV24_1 * KONV24 Crosstabulation

Count

		KONV24		Total
		,00	44,00	
DRV24_1	71,00	5	1	6
Total		5	1	6

DRV25_1 * KONV25 Crosstabulation

Count

		KONV25			Total
		-58,00	-24,00	,00	
DRV25_1	3,00	1			1
	37,00		2	1	3
	61,00			2	2
Total		1	2	3	6

DRV26_1 * KONV26 Crosstabulation

Count

		KONV26						Total
		-41,00	-13,00	-7,00	11,00	17,00	32,00	
DRV26_1	44,00	1		1		1	1	4
	50,00		1		1			2
Total		1	1	1	1	1	1	6

Tabelle A6: Kontrasturteile, identische Schätzungen und Assimilation pro Frage (konv{x}) bei allen Befragten (Alle), bei Experten (Exp.) und bei Nichtexperten (Nexp.)

Variable	Kontrasturteil			identische Schätzungen			Assimilation			Missing		
	Alle	Exp.	Nexp.	Alle	Exp.	Nexp.	alle	Exp.	Nexp.	Alle	Exp.	Nexp.
2	2	1	1	6	5	1	2	-	2	-	-	-
3	3	1	2	3	3	-	4	2	2	-	-	-
4	4	1	3	5	4	1	1	1	-	-	-	-
5	-	-	-	7	5	2	2	-	2	1	1	-
6	4	3	1	1	-	1	5	3	2	-	-	-
7	-	-	-	8	5	3	1	-	1	1	1	-
8	1	1	-	4	3	1	5	2	3	-	-	-
9	2	-	2	7	5	2	-	-	-	1	1	-
10	2	-	2	5	4	1	3	2	1	-	-	-
11	-	-	-	9	5	4	-	-	-	1	1	-
12	2	-	2	2	2	-	6	4	2	-	-	-
13	-	-	-	8	5	3	1	-	1	1	1	-
15	4	2	2	5	4	1	1	-	1	-	-	-
16	2	1	1	4	3	1	4	2	2	-	-	-
17	1	1	-	6	3	3	3	2	1	-	-	-
18	3	2	1	4	2	2	2	1	1	1	1	-
19	1	-	1	9	6	3	-	-	-	-	-	-
20	2	1	1	5	3	2	3	2	1	-	-	-
21	-	-	-	-	-	-	10	6	4	-	-	-
22	-	-	-	8	6	2	2	-	2	-	-	-
23	1	1	-	8	4	4	1	1	-	-	-	-
24	1	1	-	8	5	3	1	-	1	-	-	-
25	1	-	1	5	3	2	3	3	-	1	-	1
26	5	3	2	-	-	-	4	3	1	1	-	1
27	-	-	-	7	5	2	3	1	2	-	-	-
28	3	2	1	5	3	2	2	1	1	-	-	-
29	1	-	1	3	2	1	5	4	1	1	-	1
30	3	2	1	5	3	2	1	1	-	1	-	1
31	3	1	2	4	2	2	3	3	-	-	-	-
32	-	-	-	10	6	4	-	-	-	-	-	-
33	3	2	1	3	1	2	3	2	1	1	1	-
34	-	-	-	9	5	4	-	-	-	1	1	-
35	2	1	1	4	2	2	4	3	1	-	-	-
36	-	-	-	8	5	3	2	1	1	-	-	-
Summe	56	27	29	185	119		87	50	37	12	8	4

Anhang III: Instrumente

Schriftlich zu beantwortende Fragen in der ersten Welle¹⁴

¹⁴ Die Nummerierung der Indikatoren bei den gezielten Nachfragen folgt hier der im Projekt "Zukunft des Handwerks" benutzten.

Mündliche Zusatzfragen in der ersten Welle

Schriftlich zu beantwortende Fragen in der zweiten Welle einschließlich Rückmeldung der Ergebnisse aus der ersten Welle

Mündliche Zusatzfragen in der zweiten Welle